(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 December 2003 (04.12.2003)

PCT

(10) International Publication Number
WO 03/100372 A1

(51) International Patent Classification[7]: G01L 15/00, 11/02

(21) International Application Number: PCT/FI03/00400

(22) International Filing Date: 26 May 2003 (26.05.2003)

(25) Filing Language: Finnish

(26) Publication Language: English

(30) Priority Data:
20025028 29 May 2002 (29.05.2002) FI

(71) Applicant *(for all designated States except US)*: NOKIA CORPORATION [FI/FI]; Keilalahdentie 4, FIN-02150 Espoo (FI).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: KINNUNEN, Kimmo [FI/FI]; Peltokatu 23, FIN-44100 Äänekoski (FI). RANTANIVA, Mika [FI/FI]; Emännäntie 21 B 38, FIN-40740 Jyväskylä (FI). LEHTIMÄKI, Matti [FI/FI]; Kyyhkysmäki 16 C 38, FIN-02600 Espoo (FI).

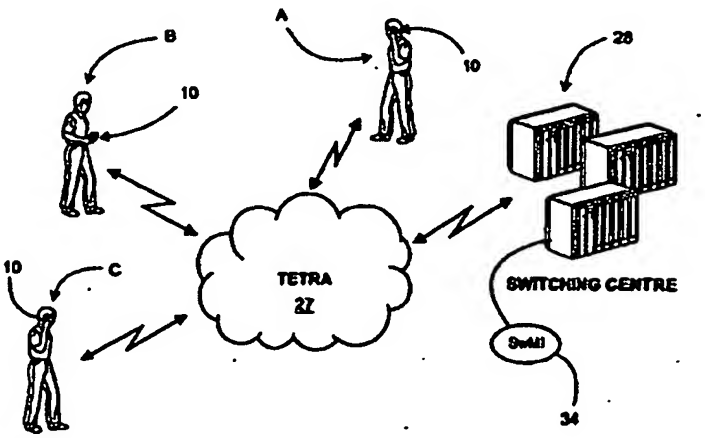
(74) Agent: KESPAT OY; P.O.Box 601, FIN-40101 Jyväskylä (FI).

(81) Designated States *(national)*: AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD IN A DIGITAL NETWORK SYSTEM FOR CONTROLLING THE TRANSMISSION OF TERMINAL EQUIPMENT



(57) Abstract: The invention concerns a method in a digital network system (27) for controlling the transmission of terminal equipment (10). Terminal equipment (10) includes a PTT (Push-to-Talk) function in order to at least activate the transmission to be carried out to the said network system, and wherein the terminal equipment (10) for voice control of the said PTT function also includes a VOX (Voice Operated transmission) feature, which can be activated/passivated and which is implemented by a VRE (Voice Recognition Engine) function (23). In the method steps are performed - the VRE function (23) is used to search for an established keyword from an audio signal (406, 407), - the established keyword is recognised from the audio signal (408), - a turn to transmit is requested from the network system (27) (409), - a turn to transmit is received from the network system (27) (412), - the transmission is connected and the granted turn to transmit is indicated (413, 414), - the transmission is carried out (415), and - the transmission is passivated (419). In the said VOX feature before the said VRE function (23) the audio signal is monitored 25 by a VAD. (Voice Activity Detection) function (22) arranged in connection with terminal equipment (10), and whereby when activating the said VOX feature (401, 402) in the terminal equipment (10) steps are performed before the said partial steps (406 - 419) - the terminal equipment's (10) incoming audio signal is processed with the VAD function (22) searching it for a signal form in accordance with an established criterion (404, 405), and - when a signal form according to the established criterion is detected in the audio signal, the said VRE function is activated to search for an established keyword (405, 406).

WO 03/100372 A1

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD IN A DIGITAL NETWORK SYSTEM FOR CONTROLLING THE TRANSMISSION OF TERMINAL EQUIPMENT

The invention concerns a method in a digital network system for controlling the transmission of terminal equipment, which terminal equipment includes a PTT (Push-to-Talk) function at least to activate the transmission to be carried out to the said network system, and wherein the terminal equipment also for voice control of the said PTT function includes a VOX (Voice Operated transmission) function, which is activated/passivated and which is implemented by a VRE (Voice Recognition Engine) function, and in which method the following takes place by steps

- the VRE function is used to search for an established keyword from an audio signal,
- the established keyword is recognised from the audio signal,
- a turn to transmit is requested from the network system,
- a turn to transmit is received from the network system,
- the transmission is connected and the granted turn to transmit is indicated,
- the transmission is carried out, and
- the transmission is passivated.

The invention may also be applied in PoC (Push-to-talk over Cellular) speech services systems.

In digital radio network systems, such as, for example, the TETRA (TErrestrial Trunked RAdio) system, semiduplex communication represents an efficient mode of communication from the viewpoint of system capacity. Usual bottlenecks in system capacity are the limited bandwidth and the system's processing ability. In order to carry out semiduplex

communication only one downlink traffic channel is needed for the broadcasting from the base transceiver station to the terminal equipment and one uplink traffic channel is needed for the transmitting terminal equipment. Communication on the channels is by so-called broadcasts, which the TETRA switching centre transmits to all pieces of terminal equipment, even if a message is intended for one of them only. In order to arrange uplink traffic a definite method of trunking is usually required, which is used to organize the transmissions of the terminal equipment.

However, it is a requirement in semiduplex communication that there is only one transmitting party at a time in the system. This requirement is typically met with the PTT (Push-to-Talk) switch of the terminal equipment, which the user must push when wishing to transmit. Pushing of the PTT switch produces a request for a turn to transmit, based on which the trunking system of the TETRA switching centre grants one party at a time a turn to talk based on his talking turn indication algorithm.

All the parties engaged in semiduplex communication, both in group and also in direct private calls (personal semiduplex), must also in the TETRA system, which is e.g. used by the authorities, request and obtain their turns to transmit before their turn to talk. Traditionally, this has been implemented by using the PTT switch of the terminal equipment. However, this method restricts essentially the activity of the user of the terminal equipment during the communication, because he must use one hand to press the PTT switch.

There are several practical situations both in office and field conditions, where it would be practical to have both hands free. Examples of such situations are communications carried on in vehicles, such as when driving a motorbike or a car, and

further, for example, in the case of an electrician commenting on his electric installation, when the electrician needs both his hands in order to figure out the installation or for other such measures.

Furthermore, situations of a similar kind where both hands must be free also occur in connection with terminal equipment supporting the PoC (Push-to-talk over Cellular) feature/service. It is a characteristic of the PoC feature that it is implemented as a duplex radio service of a kind known as such. The user of the terminal equipment can there be in a constant connection, practically speaking, with his own group, but despite this, maintenance of the connection does not keep the transmission channel busy all the time.

When using the PoC feature, the user pushes the transmission key in the earpiece of his terminal equipment, whereupon he can immediately say the message to be transmitted. All such parties belonging to the same group as the user, who at the time of transmission are connected to the data communication network, for example, over a packet connection (such as GPRS), will hear the message. The PoC feature also supports at least two transmission modes. In the first mode, one of the parties may address a group call to the other parties, while in the second mode one of the parties may address a direct call to some other party.

In addition to the above-mentioned traffic situations, situations requiring free use of both hands when using the PoC feature may occur, for example, when playing network games. Hereby the players give comments to the other parties as the game proceeds. According to the state of the art, a manual connection must be made in order to carry out the transmission. Another problem is that the user cannot easily carry on private

communication with another certain party when a group call is going on.

The VOX function, that is, Voice Operated transmission, is a feature known from some analog PMR (Private Mobile Radio) pieces of terminal equipment used in semiduplex communication. In these, the VOX feature allows requesting a turn to talk without pushing the PTT switch manually.

The use of VAD (Voice Activity Detection) is known from the implementation of DTX (Discontinuous Transmission). Hereby, when a voice is not detected with VAD in the microphone signal during the call, the terminal equipment is not either used for transmitting whole uplink bursts corresponding with these quiet moments. The function is used to save duration of the transmission power and thus to prolong the effective operating time of the terminal equipment.

Some types of mobile station terminal equipment are nowadays equipped with a talk detection feature. In these, the user can control the terminal equipment by uttering a command he has defined, such as, for example, the name of the party to be called, "Charlie". In consequence of the command, the terminal equipment activates the subscriber identity of the party to be called (Charlie's). According to the user's choice, the subscriber identity activated by the command can be acknowledged by pushing a key, or the terminal equipment may also without any action on the part of the user call the subscriber identity of the party activated by the command.

Due to the constant consumption of current of the active audio parts, such as, for example, the processor processing audio data, it is very disadvantageous especially in mobile terminal equipment to implement such a VOX function based on a program-

based solution, which constantly meets the incoming audio signal and detects talk or individual words in this.

Many methods of implementing VOX have been proposed, but these have usually been based on hardware-level solutions, such as, for example, integrated additional VOX circuits or separate circuit diagrams. Drawbacks of solutions of this kind are increased component costs, the additional space needed by components and obviously also the increased consumption of current of the terminal equipment's bigger standby space. Software-based implementations are also known, such as, for example, the above-mentioned constant audio monitoring. The increased current consumption also restricts the use of these in mobile terminal equipment. These solutions may, however, be functioning e.g. in car installation series, where current consumption is not a major problem as such.

As regards the state of the art, reference is made to PCT publication WO-96/11529 and to US patent 5,912,882. Publication WO 96/11529 presents activation of the transmission of a radio telephone by using the voice recognition function. Here the terminal equipment performs continuous recognition of keywords on the audio data. However, power is consumed considerably in the use of the voice recognition (VRE) function based on recognition of words for activating connection of the transmission, which is a real problem, especially with pieces of mobile terminal equipment.

Publication US-5,912,882 presents implementation of a private communication system in a PSTN network. This includes a mention of activation of the PTT facility by voice control. However, this is not a genuine digital network system, but the signal undergoes DA conversion when moving from a wireless network (CDMA) to a telephone network (PSTN). However, activation of

the PTT function by voice recognition will not function at all in practical situations, because in principle the transmission is activated by every audio signal recognisable as a voice or generally, for example, as talk. In addition, passivation of the PTT is performed by detecting a pause, the duration of which is established in advance.

All things considered, it is difficult to bring about a functioning and especially a reliable and efficient VOX function with state-of-the-art solutions, particularly in mobile terminal equipment, for example, especially in a digital trunking system, where the terminal equipment must make a request to the trunking system for a turn to talk.

It is the purpose of this invention to bring about an essentially more advantageous, more user-friendly and more reliable method for controlling transmission of the terminal equipment in a digital network system. The characteristic features of the method according to the invention are presented in claim 1.

The method according to the invention makes possible implementation of the VOX feature in its simplest form in every piece of terminal equipment with existing VAD (Voice Activity Detection) and VRE (Voice Recognition Engine) algorithms, which are preferably used in accordance with the method of the invention in detection of the audio signal and in searching this for one or more keywords. The VRE function can be implemented simply with audio DSP (Digital Signal Processing) algorithms and it may be used for detecting in the audio signal a request for a turn to talk and also generally key words activating the transmission, depending on the network system being used.

Activation of the feature may be done with a special UI (User Interface) concept, thus allowing its flexible on/off switching. In practice, this means that the user of the terminal equipment must first activate the VOX feature in some way, whereupon the feature is active, for example, for an established period of time, a logical sequence or according to choices made by the user in the UI.

The method according to the invention improves essentially the usability of the terminal equipment in semiduplex traffic. An advantage is attained in trunking systems, such as the TETRA. With the feature in question advantages are also attained in PoC (Push-to-talk over Cellular) group communication, which is one embodiment of the VoIP (Voice over Internet Protocol) professional talk services designed for All-IP-based systems. One of their objectives is to control the talk for transfer as IP packets, for example, through the GPRS system.

When implemented entirely on a software basis without any additional equipment or components installed in the terminal equipment, the VOX feature as a combination of VAD and VRE functions significantly reduces variable costs, reduces the size of the printed circuit board of the terminal equipment and reduces the basic current consumption in particular. When implemented in accordance with the method, the feature can be implemented advantageously on existing known product platforms, because their audio parts usually include the required VAD and VRE functions. The software-based solution and the user interface concept give many possibilities of configuring the settings relating to the function, such as, for example, its ON/OFF feature and activation and passivation settings according to the needs of the users.

According to one embodiment, the method according to the invention may also be used, for example, in the above-mentioned PoC group communication. Hereby the concept may be different from trunking systems, for example, as regards the types of talk and allocation of turns to talk. In PoC group communication, the method according to the invention may be utilised as an additional form of application, besides the said activation of transmission, for a combined choice of recipient.

Other additional advantages achieved with the method according to the invention emerge from the specification part, while the characteristic features emerge from the appended claims.

The method according to the invention, which is not limited to the embodiments to be presented hereinafter, will be described in greater detail by referring to the appended figures, wherein

Figure 1	shows an example of the functional parts of a terminal equipment,
Figure 2	shows an example of an application of the method according to the invention,
Figures 3a and b	are flow diagrams showing an example of an embodiment of the method according to the invention,
Figures 4a and b	are flow diagrams showing an example of another embodiment of the method according to the invention, and
Figure 5	shows another example of an application of the method according to the invention.

Figure 1 shows an example of the functional parts of a digital terminal equipment 10 implementing the method according to the invention. In connection with the processor unit 18 of the

terminal equipment 10 a transmitter-receiver circuit 19 is arranged, in connection with which an antenna 25 is connected, among other things, to carry out and receive transmission. Furthermore, in connection with processor unit 18 there are keyboard 11 of the terminal equipment 10, navigation and selecting keys 15 and switches as well as a possible SIM (Subscriber Identity Module) card 16. Among other things, a PTT (Push-to-Talk) switch 26 controlling a possibly occurring request for a turn to transmit and controlling the transmission also belongs to the said switches.

The terminal equipment 10 may include an LCD display 21, which is arranged in connection with a display controller 13, which is also arranged in connection with processor unit 18. Furthermore, in connection with processor unit 18 are arranged RAM memory 17a and up-datable ROM memory 17b as well as audio part 14, in connection with which are arranged loudspeaker and microphone means 12, 20a of a kind known as such as well as a possible buzzer 20b. It should be noticed that the functional parts of the terminal equipment 10 shown in Figure 1 are shown in quite a rough manner by way of example. Terminal equipment 10 may be implemented in many different ways, for example, depending on its type, but these are obvious to the man skilled in the art.

It is essential for the method according to the invention that, for example, in the audio part 14 of terminal equipment 10 an algorithm module 22, that is, voice detection, is arranged as a software sub-section implementing the VAD (Voice Activity Detection) function. According to a more advanced embodiment, the functionality of the audio part 14 includes as a sub-section, besides the VAD module 22, also a DSP module, which includes a VRE (Voice Recognition Engine) function 23, that is, voice recognition.

In the following some advantageous embodiment of the invention will be described with reference made to Figures 2, 3a and 3b. Figure 2 is a schematic view of an application of the method according to the invention. Users A, B and C, who may be, for example, police officers on patrol in the field or representatives of some other such authority, business enterprise or public transport department, have pieces of terminal equipment 10 according to the functionality shown in Figure 1. According to one embodiment, the pieces of terminal equipment 10 are intended to operate in a network system based on a digital trunking system, such as in the TETRA (TErrestrial Trunking RAdio) 27. It is typical of the Trunking system that when a terminal equipment 10 asks the trunking system for a turn to transmit, the system's SwMI (Switching and Management Infrastructure) will distribute turns to transmit according to established criteria. Such criteria may be, for example, the requesting order, the priority level of the users A, B, C and the active type of transmission of their terminal equipment 10 (for example, an emergency call vs. an ordinary turn to talk).

Figures 3a and 3b are flow diagrams showing an advantageous embodiment of the method according to the invention in a trunking system. Users A, B, C activate the VOX feature, for example, manually, from the user interfaces UI of their TETRA terminal equipment 10 with the ON/OFF setting (301). After this measure, the terminal equipment 10 activates a group message transmission by a brief push on the PTT switch (duration < 500 ms) (302). Upon activation of the VOX feature, a signal tone or other such notification, such as a signal light, is given, for example, with buzzer 20b of the terminal equipment (303).

When the VOX feature implemented according to the method of the invention is active, the audio path is kept open all the time.

The audio signal arriving through microphone 20a is processed without interruption in a manner known as such with the VAD algorithm (304), which is used to search the audio signal for a signal form according to the established criterion, such as, for example, possible talk of the user of the terminal equipment (305). If needed, the sensitivity of the microphone 20a and the VAD module 22 can be adjusted, in order to avoid, for example, any false transmissions connected by strong background sounds. According to the method of the invention, the VAD function 22 is used to look for the initial point of talk in the audio signal arriving by way of the microphone 20a. With the VAD algorithm fitted in connection with the VAD module 22 any rise of the signal level is detected in the audio signal arriving through microphone 20a, which rise may be talk. It is not possible with the VAD function 22 to distinguish talk or individual words from the sound.

In this first embodiment based solely on the VAD function 22, the user's A, B, C first word, with which the user can activate a request for a turn to transmit, must be something else than real talk intended for transmission. Before his utterance to be transmitted, user A, B, C must utter, for example, the word "VOX" or any other word or sound. Hereby VAD 22 detects a possible transmission and transmits a request for a turn to transmit to the network system's switching centre 28 (306). The SwMI arranged in switching centre 28 processes the request for a turn to transmit (307) and if at that time there is no traffic in the group formed by users A, B, C, SwMI will grant a turn to transmit to the requesting terminal equipment 10, usually almost immediately (308). If there is much traffic in the group, then the users have to wait for their turn to transmit, depending e.g. on the priority level of the user A, B, C sending the request. Terminal equipment 10 receives a

permission to transmit (309), and the following partial steps (310 - 317) will be explained in greater detail hereinafter.

In an embodiment based on the VAD function 22, where the user A, B, C utters the command "VOX" activating the VOX feature and then immediately the message he wishes to be transmitted, words may be left out from the beginning of the message. A way to avoid this is to reserve more memory space in order to buffer microphone talk. However, in this case longer talk delays will result, which may be no less than tens of milliseconds. Transmissions activated by strong background sounds are a significant drawback in solutions based on the VAD function only.

Another more advantageous way of implementing the method according to the invention is shown in Figures 4a and 4b. This uses the VAD function 22 presented above, and in connection with this a VRE function 23, that is, word recognition. Users A, B, C activate the VOX feature, for example, manually from the user interfaces UI of their pieces of TETRA terminal equipment 10 with the ON/OFF setting (401). After this action terminal equipment 10 activates a group message transmission by a brief push on the PTT switch (duration < 500 ms) (402). Upon activation of the VOX feature, a signal sound or other such notification, such as a cue light signal (403), is given using, for example, the buzzer 20b of the terminal equipment 10.

When the VOX feature implemented in accordance with the method of the invention is active, the audio path is kept open all the time. The audio signal arriving through microphone 20a is processed without interruption by a VAD algorithm in a manner known as such (404), which algorithm is used to search for a signal format according to the established criterion, such as, for example, possible talk of the user of the terminal

equipment (405). When required, the sensitivity of microphone 20a and VAD module 22 can be adjusted in order to avoid faulty transmissions turned on by, for example, strong background noises. Thus, according to the method of the invention, the VAD function 22 is used to search for the starting point of talk in the audio signal arriving through microphone 20a. The VAD algorithm adapted in connection with VAD module 22 is used to detect a raise of the signal level in the audio signal coming in through microphone 20a, which raise may thus be talk. The VAD function 22 cannot be used for distinguishing talk or individual words in the sound.

When the VAD function 22 detects for the first time in the audio signal (1°) a signal possibly intended into the microphone 20a by user A, B, C, the voice recognition function VRE 23 of the terminal equipment 10 is activated (406).

In voice recognition 23, a search is made in the talk coming in through microphone 20a for e.g. an utterance of "VOX" or for some other essentially predetermined key word (408). In case the established key word is not found within an established period of time, the procedure may return, for example, to step (405) to find out whether there is any such signal on the audio path that might be understood as a voice. If there is, then the procedure moves on directly to step (407) along route (2°).

After the voice recognition 23 has found the correct key word, terminal equipment 10 will send a request for a turn to transmit (409) to the SwMI 34 of the trunking system, corresponding to the pushing of PTT switch 26 to the bottom, as is done nowadays.

SwMI 34 processes the requests for turns to transmit (410) and grants it in a sequence to the requesting terminal equipment 10

(411). When the terminal equipment 10 has received the granted turn to transmit (412) from SwMI 34, the transmission is turned on (413) and it is indicated, for example, with a TX Granted tone (414). User A, B, C dictates the message to be transmitted into microphone 20a and terminal equipment 10 transmits it to the data communication network 28 in a known manner (415).

According to one embodiment of the invention, passivation of the transmission may be detected in such a way that the VAD algorithm 22 is used to process the audio signal during the transmission (313), and if a sufficiently long pause, for example, of a length set in advance (for example, 1 - 2 seconds) (314) is detected in the talk, the transmission is passivated in a corresponding manner as when releasing the PTT switch 26 (316). Then the procedures goes back to step (304), depending, for example, on the user's actions or on the settings of the VOX feature (317).

One or more special key words identifiable with the VRE function 23 constitute a more advanced embodiment for controlling the transmission. Hereby the audio signal is processed with the VAD or VRE function 23 (416) during the transmission. In the processing a search is made in the audio signal for an established ending criterion, which may be, for example, a key word (417). Another example of such an ending criterion is a pause of an established length in the talk, because it is always possible that voice recognition based on probability calculation can fail in some way. When an established keyword or a pause of an established length is found, passivation of the transmission is indicated (418) and the VRE and transmission are passivated (419). The procedure can then move on to step (404) (420).

By using keywords the users A, B and C can control when talk is being transmitted to the network system 27 and when it is not transmitted. An example of such use of a keyword could be "Vox (pause) assisting forces are needed here, over!". Hereby the recipients hear the phrase "Assisting forces are needed here, over!" Now, besides the word Vox, the word over is also set in the database dB arranged in connection with voice recognition 23. Database dB may be stored, for example, in the memory means 17a of the terminal equipment 10. When the VRE function 23 finds the word over in the talk signal, the conclusion can be drawn that the intention is to end the transmission.

The audio path may be kept open for the VOX feature during a time determined by the user or until an active group call is ended. Thereupon the VAD and VRE functions are closed in order to minimize the consumption of power.

The users A, B, C may carry out the passivation of the VOX feature, for example, by briefly pushing the PTT switch 26, whereby the feature is immediately passivated. This, too, is indicated to the user A, B, C, for example, with a tone signal or in some other suitable manner.

When needed, the VOX feature can also be temporarily cancelled. According to an advantageous embodiment, the users A, B, C may carry out the cancellation by keeping the PTT switch 26 pushed down for a long time, whereby transmission performed with the PTT switch 26 may be used instead of the VOX feature. After the transmission, the PTT switch 26 is released in a known manner, whereby the VOX feature according to the invention is once again active.

Users A, B, C may store keywords in database dB and program terminal equipment 10 within the limits set by the memory

capacity and by the voice recognition 23. When programming keywords, the user A, B, C of terminal equipment 10 teaches the voice recognition and establishes functions corresponding to the commands he has taught. The manner of implementation may be dependent on or independent of the speaker.

At algorithm level, implementation of the method according to the invention can be arranged as regards the VAD function 22 to take place, for example, at time level. Hereby a rise of the audio signal is detected, which rise must be sufficiently distinct. It is also possible to utilise recognition of the talk spectrum at frequency level. Hereby the audio signal must resemble talk, the signal of which is usually in a range of 100 Hz - 1,5 kHz. Hereby one significant criterion as regards functionality is to distinguish talk from background noise in the signal.

Figure 5 shows another application example, wherein the method according to the invention may also be used. Here the network system 32.1, 32.2, which supports, for example, the GPRS transmission mode, is connected in connection with All-IP infrastructure 31.1, 31.2, 33. Hereby the terminal equipment 10' supports, for example, the PoC group communication feature/service. Activation of the VOX feature of the terminal equipment 10' is carried out, for example, with a switch reserved for this purpose. It is possible also in PoC group communication to implement the method according to the invention in at least the two ways presented above (VAD, VAD & VRE).

In the first way of implementation, the terminal equipment 10' equipped with the PoC function is arranged in a special HF (hands free) mode. Hereby, when the VOX feature implemented in accordance with the method of the invention is activated and

the user A', B', C' says something, terminal equipment 10' will always transmit a PoC talk message packet. Buffering of packets and timing/sequencing of transmissions to recipients is controlled with the PoC server 31.1, 31.2. Recognition of the transmission may preferably be implemented with a VAD module of a basic model, which detects starting and ending points of talk in a signal possibly interpreted as talk, and based on these the transmission is controlled instead of pushing and releasing a tangent.

In the second way of implementation, implementation of the VOX feature in connection with the PoC function is based both on the VAD function and on the VRE function in the manner described above. In this case, the terminal equipment 10' capable of the PoC function carries out a specific HF (hands free) tangent keyword mode. Hereby terminal equipment 10' always transmits in the PoC function a talk message packet, when a person A', B', C' utters a password and then a sentence. This may also be implemented with the VAD and VRE modules of a basic model presented above, wherein the VAD module detects starting and ending points of a sentence and the VRE module recognizes a keyword and the transmission is controlled not by pushing/releasing a tangent but according to starting and ending points of a sentence detected by the VAD module.

In another advantageous embodiment, the user A', B', C' may store several keywords in the PoC terminal equipment 10'. Hereby it is possible for the user A', B', C' to choose such individual users from his group, to whom he addresses the transmission just by uttering, for example, the keyword stored as the identifier corresponding to the user intended to be the recipient. In this way the user may transmit private messages directly only to this certain user of his choice. The feature of the described kind can of course also be activated by hand

as a menu selection, but in certain conditions it is more natural to do this by talking.

Furthermore, according to an embodiment, the user may use a keyword consisting of two parts, which improves the distinguishing ability of the method. For example, when used as a keyword, "chat Jill" is a better combination as a keyword than just "Jill". The word "group", for example, may be stored as a keyword referring to the whole group. Different combinations may preferably be used in the method. Such combinations may be, for example, pushing a tangent when a group call is active and then uttering a keyword, such as a name, in order to choose the recipient of the transmission.

When using VAD and VRE modules in the PoC system, a non-standard additional field is added to the IP packet used in the system (RTP (Realtime Transport Protocol) packets are typically used). The additional field is noticed by the PoC server 31.1, which relays the message only to those recipients, who are mentioned in the additional field.

If the VRE module finds the receiving party in its database, a confirmation of an established form is given, which indicates a successful choice of voice. The confirmation may be, for example, a short beep sound or a repetition of the keyword to the user. After the confirmation is heard (or even before that, whereby the confirmation may also be given after the end of the sentence to be transmitted) the user may dictate the message he wishes to transmit.

Especially saving in the power consumption of the terminal equipment is achieved with the method according to the invention. For example, in a noisy environment a terminal equipment using recognition based only on keywords must

constantly process the signal on the audio path, which is not necessarily even talk. As in the method according to the invention this essentially constant process of keyword identification is not performed until such sound is detected on the audio path, which is in a frequency range preferably of talk form, significant saving in the basic power consumption is hereby achieved.

It should be understood that the above specification and the figures relating to it are only intended to illustrate the method according to the present invention. The procedural implementation of the method can be in numerous different ways, which are obvious to the man skilled in the art. Thus, the invention is not limited only to the embodiments presented in the foregoing or to those defined in the claims, but many such variations and modifications of the invention will be obvious to the man skilled in the art, which are possible within the scope of the inventive idea defined in the appended claims.

CLAIMS

1. Method in a digital network system (27) for controlling the transmission of terminal equipment (10), which terminal equipment (10) includes a PTT (Push-to-Talk) function in order to at least activate the transmission to be carried out to the said network system, and wherein the terminal equipment (10) for voice control of the said PTT function also includes a VOX (Voice Operated transmission) feature, which can be activated/passivated and which is implemented by a VRE (Voice Recognition Engine) function (23), and in which method steps are performed

- the VRE function (23) is used to search for an established keyword from an audio signal (406, 407),
- the established keyword is recognised from the audio signal (408),
- a turn to transmit is requested from the network system (27) (409),
- a turn to transmit is received from the network system (27) (412),
- the transmission is connected and the granted turn to transmit is indicated (413, 414),
- the transmission is carried out (415), and
- the transmission is passivated (419),

characterized in that in the said VOX feature before the said VRE function (23) the audio signal is monitored by a VAD (Voice Activity Detection) function (22) arranged in connection with terminal equipment (10), and whereby when activating the said VOX feature (401, 402) in the terminal equipment (10) steps are performed before the said partial steps (406-419)

- the terminal equipment's (10) incoming audio signal is processed with the VAD function (22) searching it for a signal form in accordance with an established criterion (404, 405), and

- when a signal form according to the established criterion is detected in the audio signal, the said VRE function is activated to search for an established keyword (405, 406).

2. Method according to claim 1 - 2, characterized in that
- the audio signal is processed with the VAD function (22) during the transmission (416),
- the audio signal is searched for a pause of an established length (417),
- a pause of the established length is found in the audio signal, whereby the signal established to indicate ending of the transmission is indicated (418) and the transmission is passivated (419).

3. Method according to claim 1 - 3, characterized in that
- the audio signal is processed with the VRE function (23) during the transmission (416),
- the audio signal is searched for the established ending criterion (417),
- the established ending criterion is found in the audio signal, whereby the signal established to indicate ending of the transmission is indicated (418), and the transmission is passivated (419).

4. Method according to any one of claims 1 - 3, characterized in that the VOX feature is turned on for an established period of time or until the active group call ends, whereupon the VAD and VRE functions (22, 23) are passivated.

5. Method according to any one of claims 1 - 4, characterized in that the VOX feature can be temporarily cancelled with an established measure.

6. Method according to any one of claims 1 - 5, characterized in that for the VRE function (23) a special database (dB) is arranged in memory means (17a) of the terminal equipment, in which database the user stores keywords to activate and passivate the transmission.

7. Method in a digital network system (32.1, 32.2) for controlling the transmission of terminal equipment (10'), wherein the said network system (32.1, 32.2) is arranged in connection with an All-IP infrastructure (31.1, 31.2, 33) equipped with a server, and the said terminal equipment (10') is arranged to support the PoC (Push-to-talk over Cellular) feature/service and wherein the terminal equipment (10') includes a PTT (Push-to-Talk) function to at least activate the transmission to be carried out to the said network system, and wherein the terminal equipment (10') for voice control of the said PTT function also includes a VOX (Voice Operated transmission) feature, which can be activated/passivated and which is implemented by a VRE (Voice Recognition Engine) function (23), and in which method the following steps are performed while carrying out the transmission with the PTT function

- the VRE function (23) is used to search for an established keyword from the audio signal (406, 407),
- the established keyword is recognised from the audio signal (408),
- the transmission is activated (415) and
- the transmission is passivated (419),

characterized in that in the said VOX feature before the said VRE function (23) the audio signal is monitored by a VAD (Voice Activity Detection) function (22) arranged in connection with the terminal equipment (10'), and whereby when activating the said VOX feature (401, 402) at the terminal equipment (10') the

following steps are performed before the said partial steps (406 - 419)

- the terminal equipment's incoming audio signal is processed with the VAD function (22) searching it for a signal form in accordance with an established criterion (404, 405), and
- when a signal form in accordance with the established criterion is found in the audio signal, the transmission of the terminal equipment (10) is activated carrying out the said partial steps (406 - 419).

8. Method according to claim 7, characterized in that keywords are used, besides to activate the transmission, to choose the recipient (A', B', C', D') of the transmission.

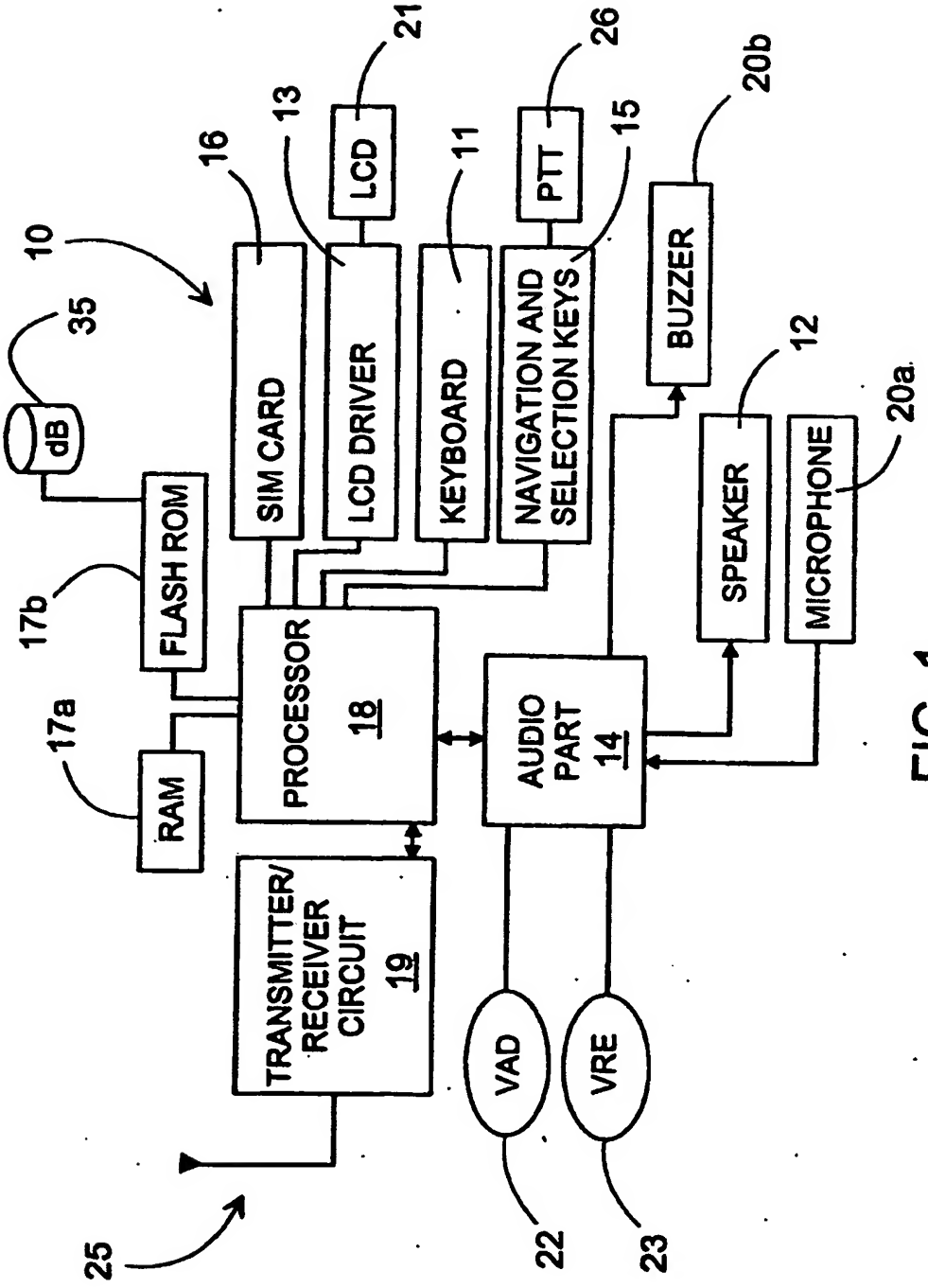

10
35
dB
17b
FLASH ROM
17a
RAM
16
SIM CARD
13
LCD DRIVER
21
LCD
11
KEYBOARD
26
PTT
15
NAVIGATION AND SELECTION KEYS
20b
BUZZER
12
SPEAKER
20a
MICROPHONE
PROCESSOR 18
AUDIO PART 14
TRANSMITTER/ RECEIVER CIRCUIT 19
VAD
VRE
22
23
25

FIG. 1

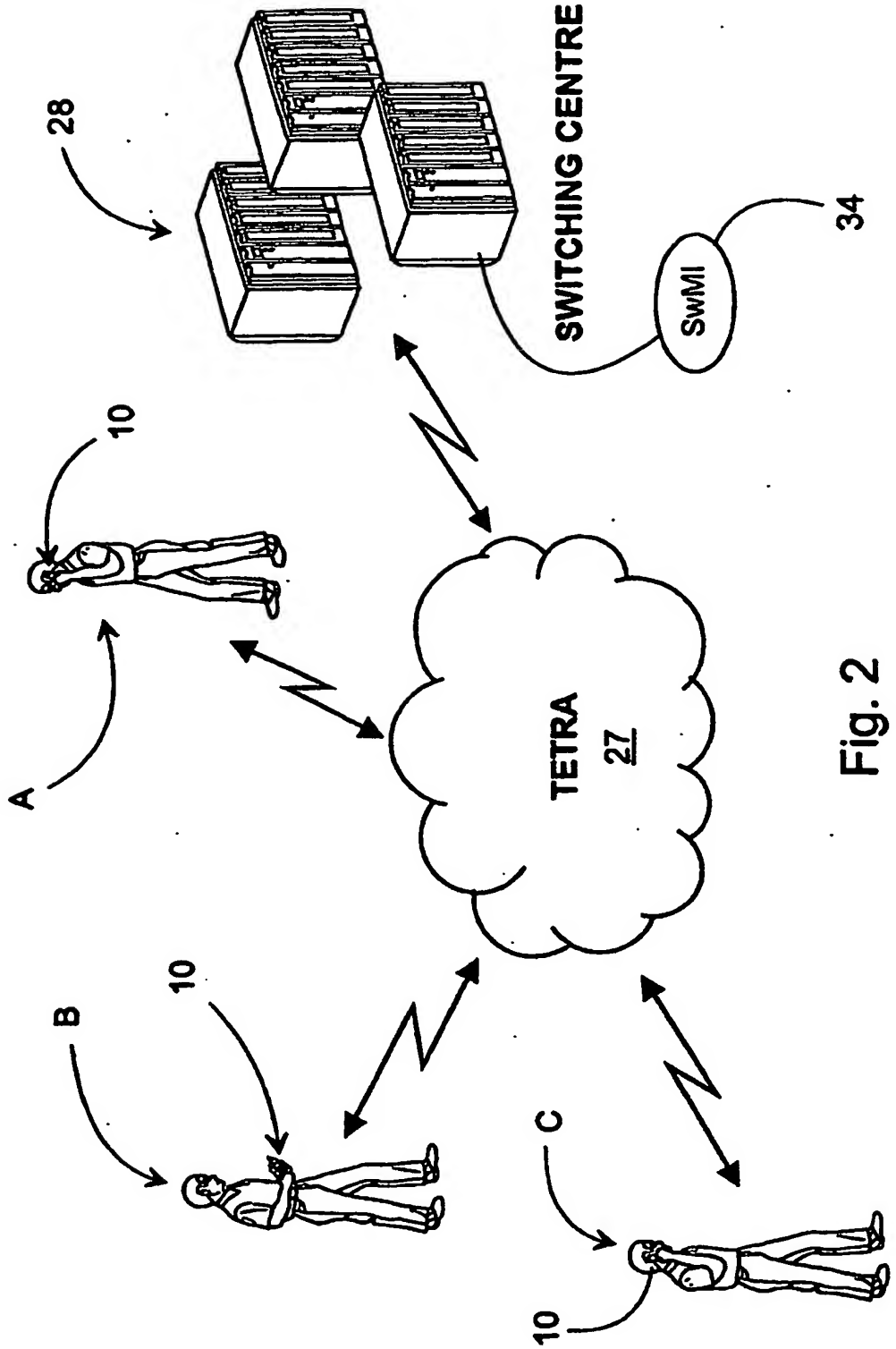


Fig. 2

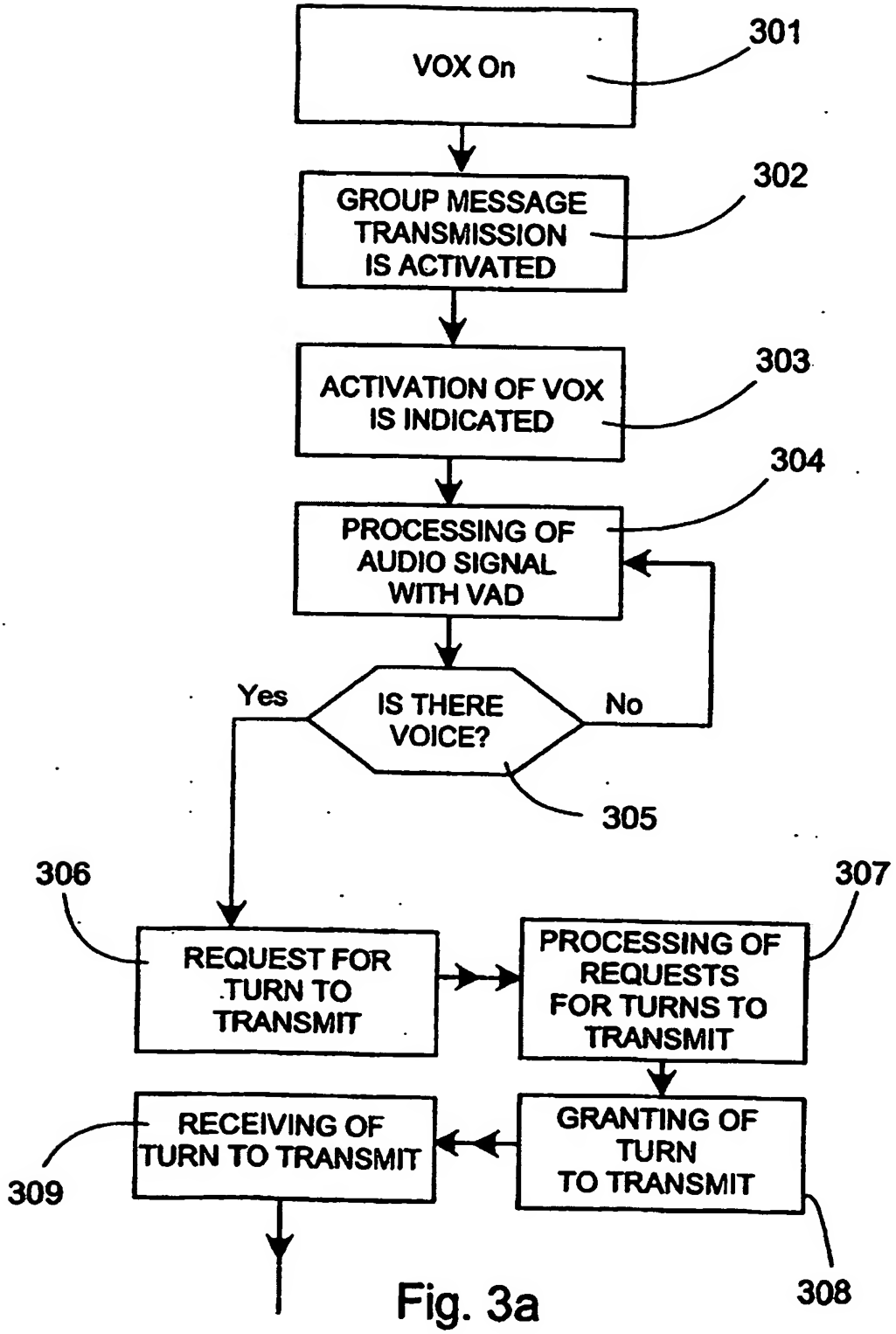

VOX On
301
GROUP MESSAGE TRANSMISSION IS ACTIVATED
302
ACTIVATION OF VOX IS INDICATED
303
PROCESSING OF AUDIO SIGNAL WITH VAD
304
IS THERE VOICE?
Yes
No
305
306
REQUEST FOR TURN TO TRANSMIT
307
PROCESSING OF REQUESTS FOR TURNS TO TRANSMIT
GRANTING OF TURN TO TRANSMIT
308
RECEIVING OF TURN TO TRANSMIT
309

Fig. 3a

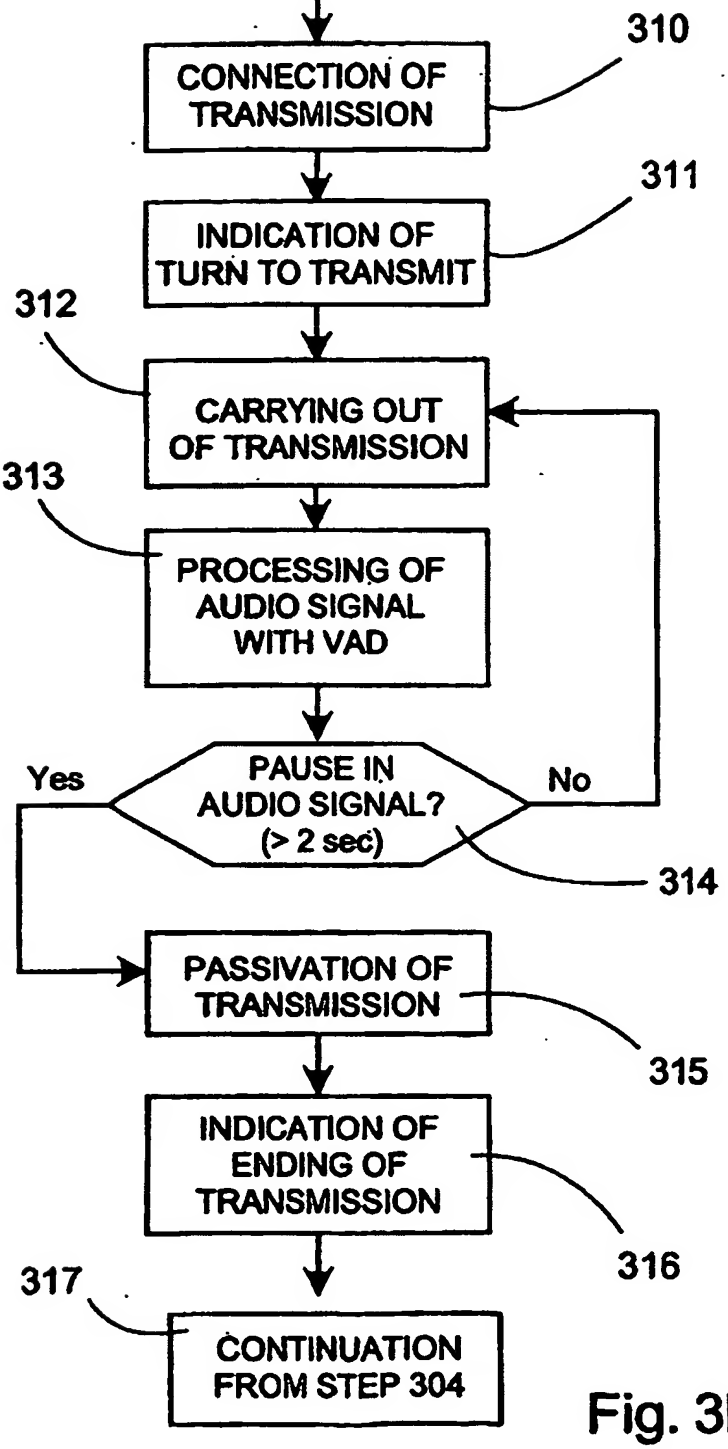

CONNECTION OF TRANSMISSION
310
INDICATION OF TURN TO TRANSMIT
311
312
CARRYING OUT OF TRANSMISSION
313
PROCESSING OF AUDIO SIGNAL WITH VAD
Yes
PAUSE IN AUDIO SIGNAL? (> 2 sec)
No
314
PASSIVATION OF TRANSMISSION
315
INDICATION OF ENDING OF TRANSMISSION
316
317
CONTINUATION FROM STEP 304

Fig. 3b

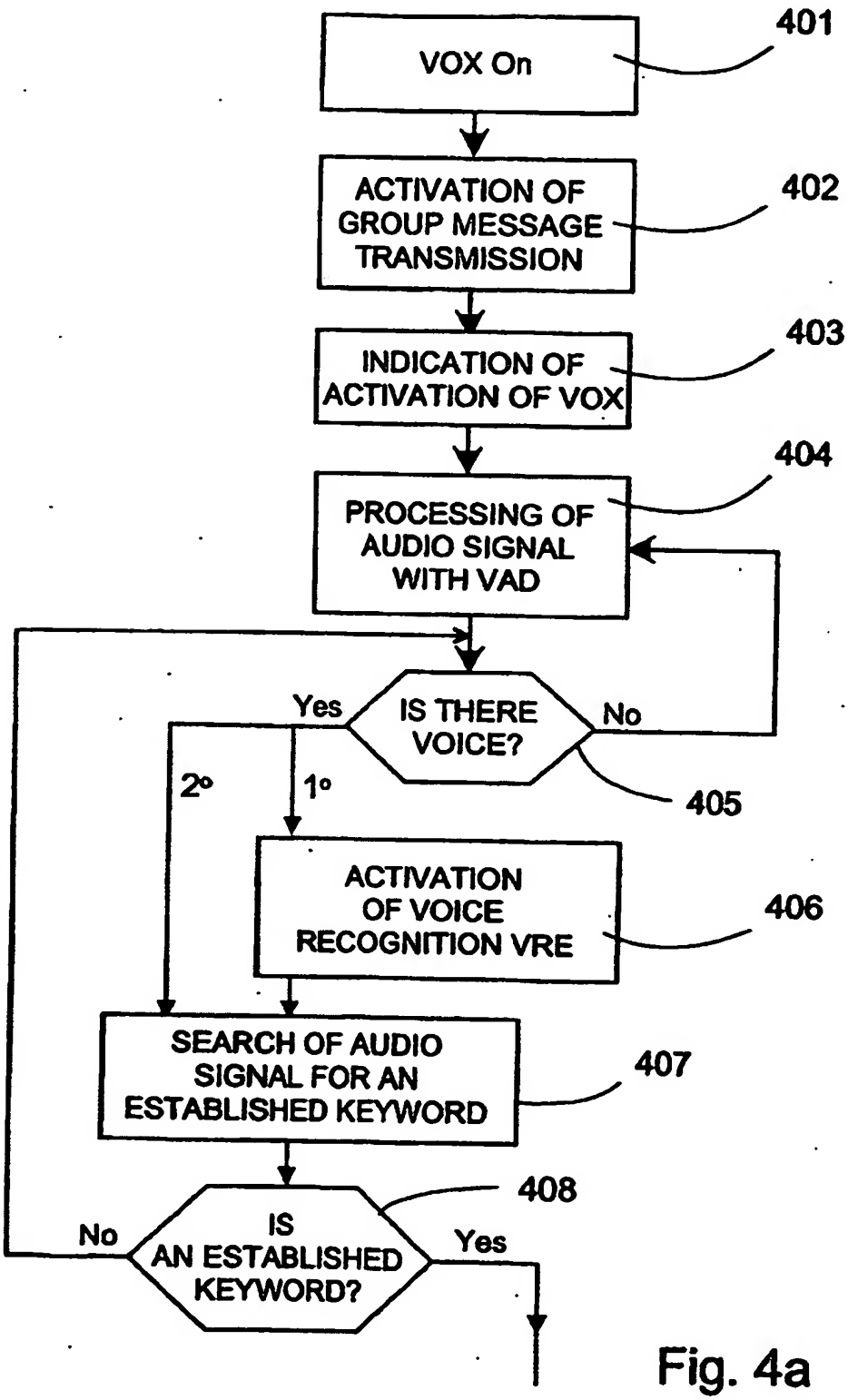

VOX On
401
ACTIVATION OF GROUP MESSAGE TRANSMISSION
402
INDICATION OF ACTIVATION OF VOX
403
PROCESSING OF AUDIO SIGNAL WITH VAD
404
IS THERE VOICE?
Yes
No
405
2°
1°
ACTIVATION OF VOICE RECOGNITION VRE
406
SEARCH OF AUDIO SIGNAL FOR AN ESTABLISHED KEYWORD
407
IS AN ESTABLISHED KEYWORD?
408
No
Yes

Fig. 4a

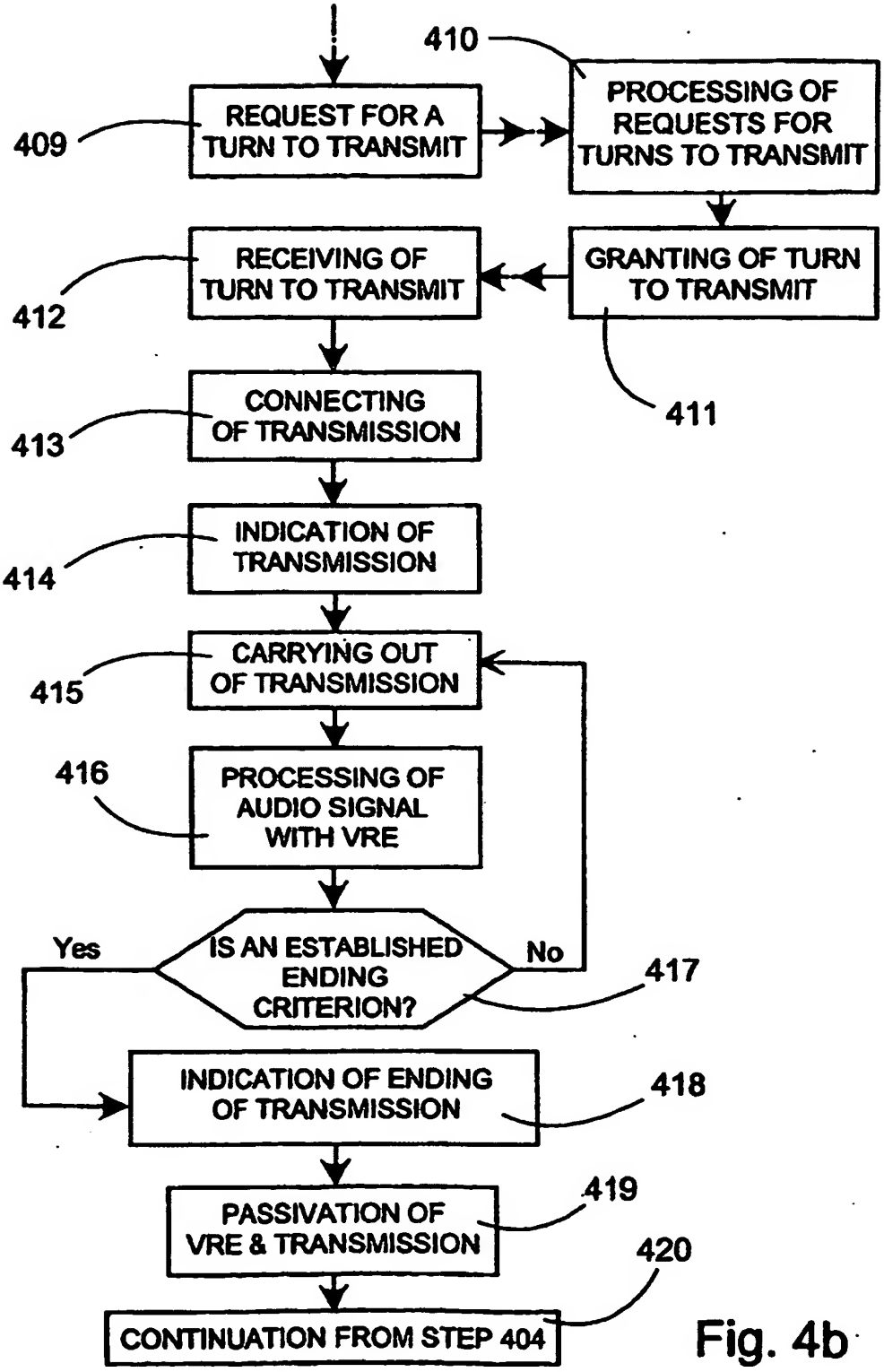

410
PROCESSING OF REQUESTS FOR TURNS TO TRANSMIT
409
REQUEST FOR A TURN TO TRANSMIT
GRANTING OF TURN TO TRANSMIT
411
412
RECEIVING OF TURN TO TRANSMIT
413
CONNECTING OF TRANSMISSION
414
INDICATION OF TRANSMISSION
415
CARRYING OUT OF TRANSMISSION
416
PROCESSING OF AUDIO SIGNAL WITH VRE
Yes
IS AN ESTABLISHED ENDING CRITERION?
No
417
INDICATION OF ENDING OF TRANSMISSION
418
PASSIVATION OF VRE & TRANSMISSION
419
420
CONTINUATION FROM STEP 404

Fig. 4b

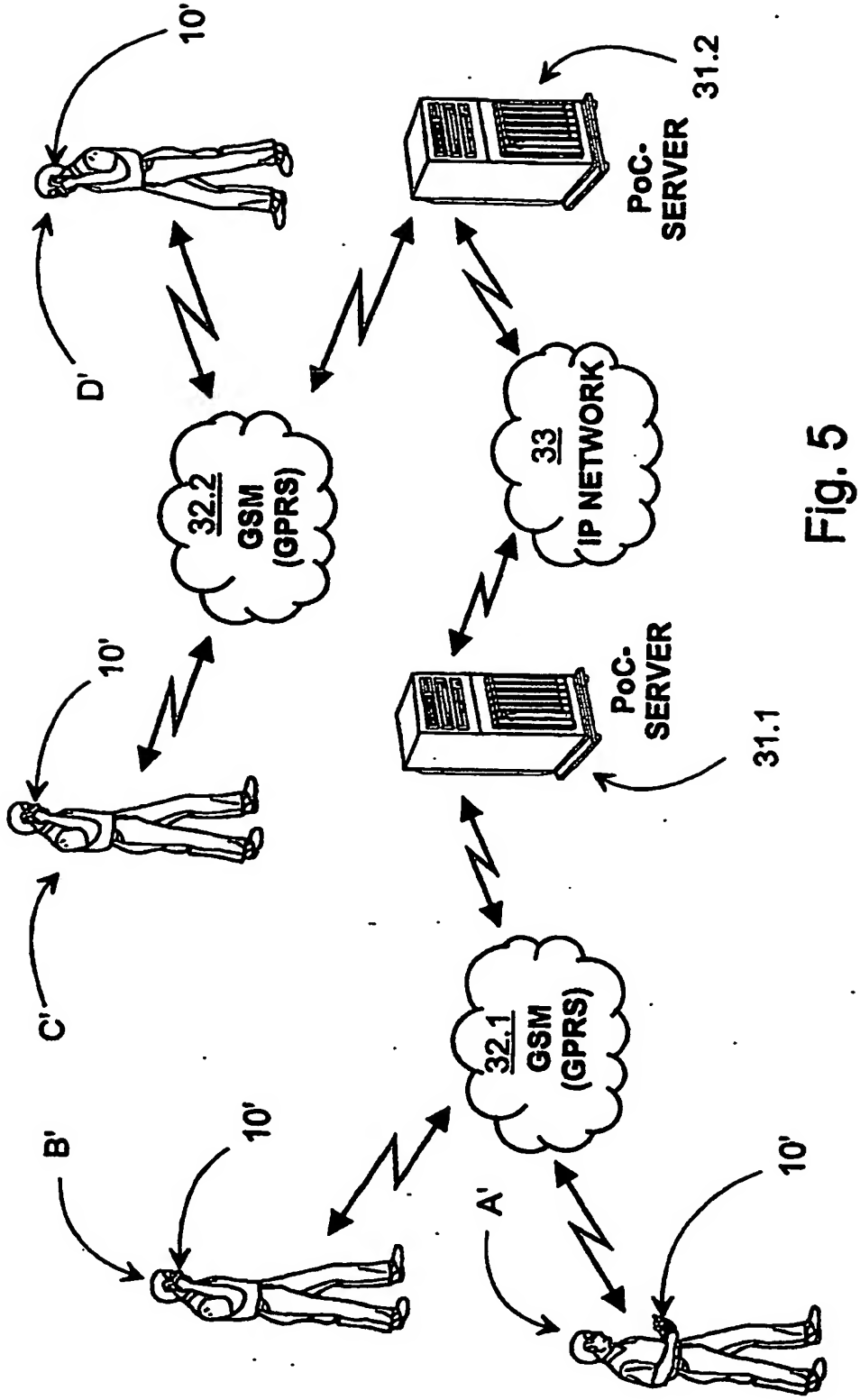

10'
D'
32.2
GSM
(GPRS)
PoC-
SERVER
31.2
33
IP NETWORK
10'
C'
PoC-
SERVER
31.1
32.1
GSM
(GPRS)
B'
10'
A'
10'

Fig. 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00400

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: G01L 15/00, G01L 11/02
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: G01L, H04M, H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-INTERNAL, WPI DATA, PAJ, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 9611529 A1 (ROTUNDA, THOMAS), 18 April 1996 (18.04.96), page 5, line 5 - line 9; page 7, line 5 - line 7; page 13, line 8 - line 22, abstract ---	1-8
Y	US 20020038211 A1 (RAJAN), 28 March 2002 (28.03.02), claim 19, abstract ---	1-8
Y	US 5912882 A (YAFUSO ET AL), 15 June 1999 (15.06.99), column 6, line 46 - line 56; column 11, line 12 - line 15 ---	1-8

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
31 July 2003	05 -08- 2003
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Peder Gjervaldsaeter/SN Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 03/00400

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1189200 A1 (PIONEER CORPORATION MEGURO-KU), 20 March 2002 (20.03.02), claim 2 --	1-8
A	WO 0131636 A2 (LERNOUT & HAUSPIE SPEECH PRODUCTS N.V), 3 May 2001 (03.05.01), abstract --	1-8
A	EP 1160768 A2 (CANON KABUSHIKI KAISHA), 5 December 2001 (05.12.01), page 18, line 48 - page 19, line 7 --	1-8
A	Sawhney, Schmandt Media Lab., MIT, Cambridge, MA; "Wearable Computers, 1998. Digest of papers. International Symposium on" Publication Date: 19-20 Oct 1998 On page(s): 108-115 INSPEC Accession Number: 6097137 see section 4.2; abstract --	1-8

Form PCT/ISA/210 (continuation of second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

Information on patent family members

29/06/03

International application No.
PCT/FI 03/00400

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9611529 A1	18/04/96	NONE	
US 20020038211 A1	28/03/02	NONE	
US 5912882 A	15/06/99	AU 716936 B	09/03/00
		AU 1848797 A	22/08/97
		BR 9707264 A	13/04/99
		CA 2244929 A	07/08/97
		CN 1214842 A	21/04/99
		EP 0878103 A	18/11/98
		FI 981689 A	18/09/98
		IL 125610 D	00/00/00
		JP 2000504182 T	04/04/00
		NO 983546 A	30/09/98
		RU 2178957 C	27/01/02
		TW 393848 B	00/00/00
		WO 9728658 A	07/08/97
		ZA 9700795 A	04/08/97
EP 1189200 A1	20/03/02	CN 1343966 A	10/04/02
		JP 2002091467 A	27/03/02
		US 2002049592 A	25/04/02
WO 0131636 A2	03/05/01	AU 1049601 A	08/05/01
EP 1160768 A2	05/12/01	GB 0018965 D	00/00/00
		JP 2002140087 A	17/05/02
		US 2002059065 A	16/05/02
		GB 0013541 D	00/00/00
		GB 0020309 D	00/00/00
		GB 0020310 D	00/00/00
		GB 0020314 D	00/00/00
		GB 2367466 A	03/04/02
		GB 2367729 A	10/04/02
		GB 2367990 A	17/04/02
		US 2002026253 A	28/02/02
		US 2002026309 A	28/02/02
		US 2002038211 A	28/03/02